

Offline 3 Report

1805012 - Aszadur Rahman Rakin

Hyperparameters

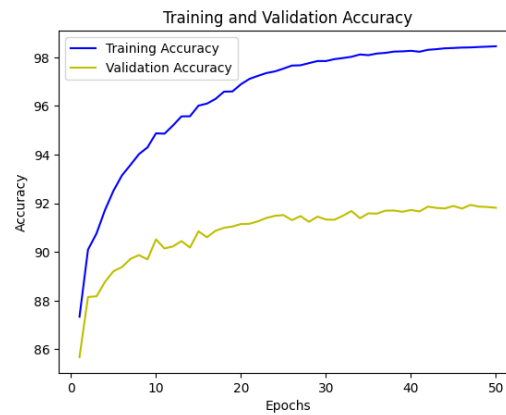
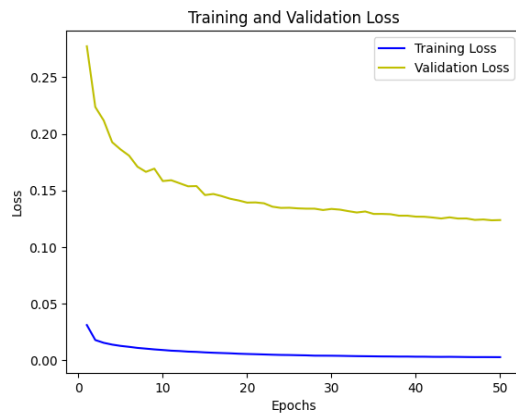
1. Learning Rate $5e-3$
2. Batch_size 1024
3. Epochs 50
4. Weight saved on Best Validation Accuracy
5. Linear LR Scheduler $(0.95^{\text{epoch}}) * \text{initial_learning_rate}$
6. Default Adam Parameters
7. Kaiming Initialization [ReLU] , Xavier Initialization [Sigmoid]

Model 1 (Best)

=====

```
Linear(784,512)
ReLU
Dropout(20.0%)
Linear(512,512)
ReLU
Dropout(10.0%)
Linear(512,26)
Softmax
```

Learning Rate: 5e-3



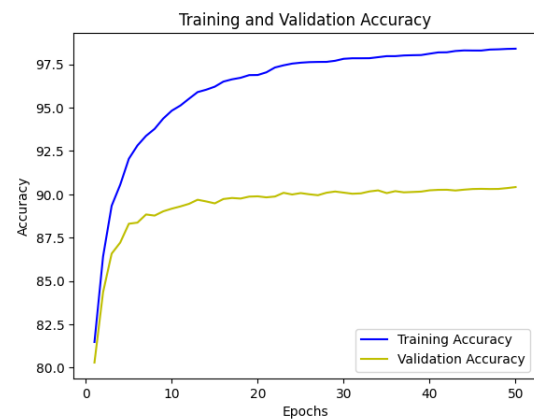
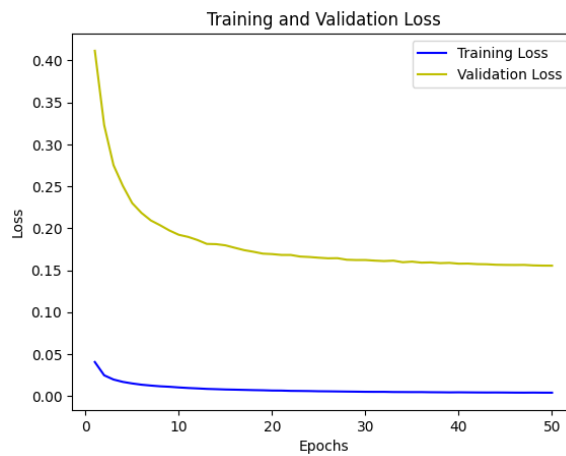
Test Accuracy: 91.73 %

Model 2

=====

```
Linear(784,1024)
Sigmoid
Dropout(30.0%)
Linear(1024,26)
Softmax
```

Learning Rate : 5e-3



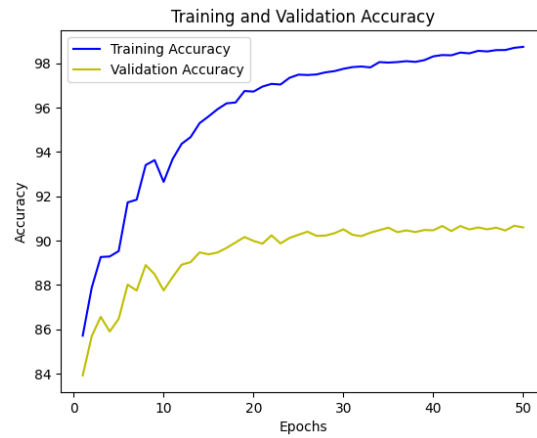
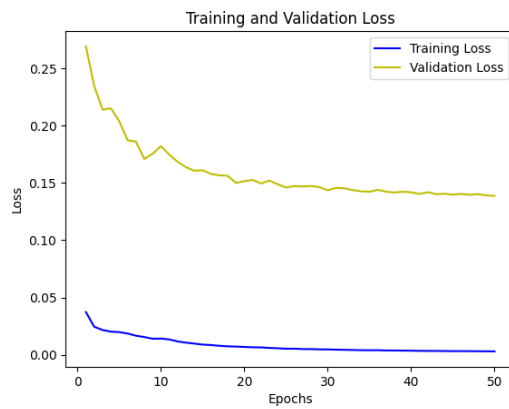
Test Accuracy: 90.23 %

Model 3

=====

```
Linear(784,1024)  
ReLU  
Dropout(30.0%)  
Linear(1024,26)  
Softmax
```

Learning Rate : 5e-3



Test Accuracy: 90.61 %